PREDICTING THE ONSET OF FILAMENTOUS BULKING IN BIOLOGICAL WASTEWATER TREATMENT SYSTEMS BY EXPLOITING IMAGE ANALYSIS INFORMATION

E.N. Banadda, R. Jenné, I.Y. Smets and J.F. Van Impe*

Chemical Engineering Department
Katholieke Universiteit Leuven, B-3001 Leuven (Belgium)
Fax: +32-16-32.29.91 e-mail: jan.vanimpe@cit.kuleuven.ac.be
*Corresponding author

Keywords: filamentous bulking, image analysis, system identification, activated sludge, ARX models.

Abstract

The performance of the activated sludge process is limited by the ability of the sedimentation tank to separate the activated sludge from the treated effluent and to concentrate it. Apart from bad operating strategies or poorly designed clarifiers, settling failures can be mainly attributed to filamentous bulking. Image analysis is a promising technique that can be used for early detection of filamentous bulking. In this work, correlations between image analysis information, i.e., the total filament length per image, the mean form factor, the mean equivalent floc diameter, the mean floc roundness, the mean floc reduced radius of gyration and classical measurements (such as the Sludge Volume Index (SVI)) have been sought and the potential of exploiting this information in ARX type black box models to predict the onset of filamentous bulking is presented.

1 Introduction

The activated sludge process is one of the most frequently used processes for the biological purification of wastewater. In the event of *bulking sludge*, there exists an imbalance between the floc forming bacteria and filamentous bacteria hence preventing formation of well settling sludge flocs [Jenkins et al., 1993]. As for now, the activated sludge process is monitored through regular microscopic observation of the sludge by the operator. In addition, sedimentation tests, such as the determination of the Sludge Volume Index[1] (SVI), are performed. However, from a pragmatic point of view, microscopic observation is time consuming and subjective (i.e., operator dependent). Image analysis, a procedure through which the microscopic images are captured and converted into digital images which can be analyzed on a computer, could provide an *objective* means to assist the operator's decision making. Furthermore, mathematical models, capable of predicting the evolution of, e.g., the SVI value, will be indispensable for the development of early warning and detection tools. Although filamentous bulking has been studied intensively during the last decades [Jenkins et al., 1993, Glasbey and Horgan, 1995, Cenens et al., 2002, Jenné et al., 2003], the phenomenon is so complex (i.e., influenced by so many different factors) that a first principles model is still lacking. The aim of this research is to investigate whether image analysis information (e.g., the total filament length, the mean form factor etc.) is correlated with classical measurements (such as the SVI) and whether this information can be exploited in ARX models for predicting the onset of filamentous bulking. Instead of developing *static* (instantaneous) correlation models as reported by da Motta and coworkers [da Motta et al., 2002], we will focus on *dynamic* ARX models.

[1] volume in milliliters occupied by 1 g of a suspension after 30 minutes settling

2 Materials and methods

Lab-scale experiments were set up that mimic large scale continuous systems, and a daily recording of the sludge characteristics, image information and settleability of the sludge was performed.

2.1 Lab-scale activated sludge system

A lab-scale activated sludge system was designed and built to simulate a full-scale installation. The laboratory set-up is a continuous type activated sludge system in a classical configuration: an aeration tank (5.5 L) followed by a sedimentation tank (3 L) and sludge recycle (Figure 1). The system was inoculated with activated sludge from a domestic wastewater treatment plant at Huldenberg (Belgium), and fed with synthetic wastewater with sodium acetate as the sole organic substrate [Houtmeyers et al., 1980], corresponding to a maximum chemical oxygen demand (COD) of 1000 mgL^{-1}. The biomass was kept as stable as possible through wastage of mixed liquor whenever necessary, yielding a sludge concentration between 2 and 3 gL^{-1}, and a sludge loading of around 0.3 g COD g MLSS^{-1} d^{-1}. Compressed air was supplied abundantly in order to meet the dissolved oxygen (DO) demands of the biomass and to ensure homogeneous mixing in the aeration tank. The DO varied between 6 and 8 mgL^{-1}. Some standard measurements, i.e., MLSS, SVI, SS and effluent COD were performed and monitored daily for a period of 100 days. The daily activated sludge composition was monitored through microscopic observation and digital image analysis.

Figure 1: Left: lab-scale activated sludge system. Right: microscope (Olympus BX51) and video camera (Sony DXC-950P) used for daily capturing of activated sludge images.

2.2 Image analysis procedure

The activated sludge images were captured using a light microscope (Olympus BX51) equipped with a 3CCD color video camera (Sony DXC-950P); this equipment is shown in Figure 1. The magnification of the microscope objective used was 10 times. A fully automatic image analysis method for recognition and characterization of both flocs and filaments in an activated sludge sample has been developed in previous work [Jenné et al., 2002, Cenens et al., 2002], and is applied to the captured images of this experiment.

2.3 Measurements

Once the objects in the image (i.e., the flocs and the filaments) are distinguished from the background, several size and shape related parameters can be computed. These parameters do not only allow the discrimination between flocs and filaments but can additionally be used for monitoring purposes to detect changes in the (settling) characteristics of the sludge.

Size measurement. The size of the sludge flocs is an important parameter with respect to the settling properties [Ganczarczyk, 1994]. The size of the flocs is expressed as the equivalent circle diameter D_{eq}, calculated from the real projected area A:

$$D_{eq} = 2\sqrt{A/\pi} \tag{1}$$

Shape measurements. It is mentioned in the literature [Eriksson and Hardin, 1984] that the shape of sludge flocs is related to the settling properties. Many shape quantifying parameters can be measured by means of image analysis [Russ, 1990, Pons et al., 1993]. Three parameters are considered in this study:

- The form factor (FF) describes the deviation of an object from a circle. It is particularly sensitive to the *roughness* of the boundaries. A circle has an FF equal to one.

$$\mathrm{FF} = 4\pi \frac{\text{area}}{\text{perimeter}^2} \tag{2}$$

- The roundness (R) is mainly influenced by the elongation of an object. It varies between 0 and 1. A circle has an R equal to one.

$$\mathrm{R} = \frac{4 \cdot \text{area}}{\pi \text{length}^2} \tag{3}$$

- The reduced radius of gyration (RG) is also influenced by the elongation of an object. A more elongated floc will have a larger RG. A circle has an RG of $\frac{\sqrt{2}}{2}$.

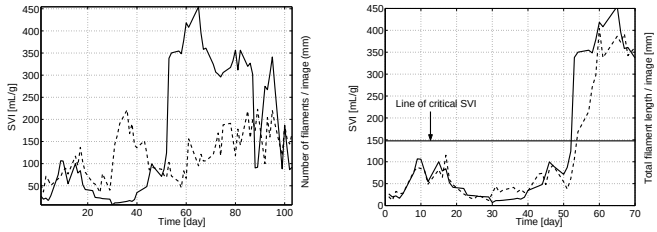
$$\mathrm{RG} = \frac{\sqrt{M_{2x} + M_{2y}}}{\frac{D_{eq}}{2}} \tag{4}$$

M_{2x} and M_{2y} are second order moments.

3 Results and discussion

3.1 Image analysis information versus SVI evolution

Correlations between image analysis information and the SVI value are sought. Figure 2 depicts the evolution of the SVI value, the number of filaments per image and the total filament length per image with respect to time.



(a) (b)

Figure 2: (a) Evolution of the SVI value (—) and the number of filaments per image (- - -). (b) Evolution of the SVI value (—) and the total filament length per image (- - -).

As can be seen from Figure 2(b), the settleability of the sludge was generally satisfactory in the period ranging from the 1st to the 52nd day: the SVI values are lower than the critical value of 150 mgL^{-1}. However, it is worth mentioning that between the 10th and the 18th day, SVI values of more than 100 mgL^{-1} were recorded. Thereafter, within the period of one week, the SVI gradually decreased to very low values (around 25 mgL^{-1}), which lasted for approximately 20 days. Then, the SVI rose again within 1 week, and on the 53rd day the critical value of 150 mgL^{-1} was exceeded, as a severe filamentous bulking event occurred. During the following 33 days, the sludge settling properties continued to be of very low quality, yielding an SVI that ranged from 300 to 450 mgL^{-1}.

Filament characteristics. Figure 2(a) illustrates that the evolution of the number of filaments per image only seems to have a relation with the SVI during the first 30 days of the experiment. The following periods of very low and very high SVI, respectively, show no instantaneous correlation between the number of filaments per image and the settleability of the sludge. This observation can be readily explained as follows.

In the period of low SVI (26th to 45th day), a large number of very short filaments was present in the sludge, without having a negative influence on the settleability. In the period of elevated SVI (after the 53rd day), the overabundance of filaments caused them to touch or overlap in the image. Under these conditions filaments could not be distinguished as individual objects, thus yielding a smaller filament number than would be expected. It is therefore advisable to consider global characteristics of filaments instead of individual ones. On the other hand, Figure 2(b) shows a clear correlation between the total filament length per image and SVI. It is to be noticed that the total filament length increased strongly on the 51st day, which is two days before the severe bulking event occurred.

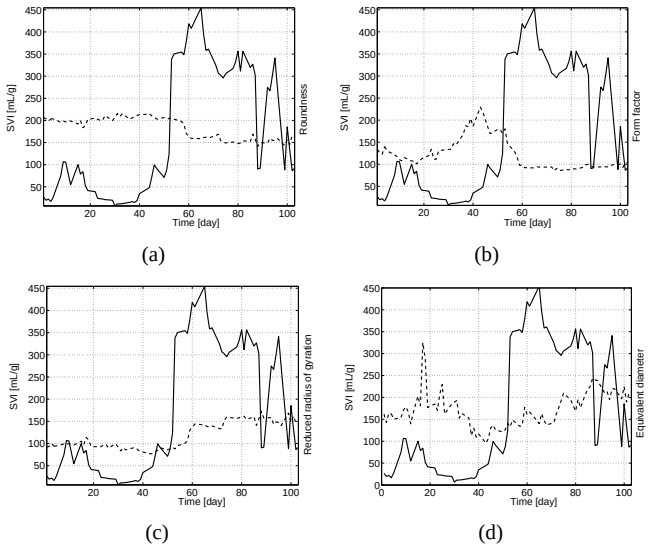


Figure 3: (a) Evolution of the SVI (—) and (a) the mean floc roundness R (- - -) (b) the mean floc form factor FF (- - -) (c) the mean floc reduced radius of gyration RG (- - -) (d) the mean floc equivalent diameter D_{eq} (- - -).

Floc characteristics. It is equally important to consider the effect of floc characteristics on sludge settleability, because sludge settling properties are not only associated with the amount of filaments. In contrast with filaments, flocs can be considered as individual objects, and several authors already associated individual floc shape with settling problems [da Motta et al., 2002, Jenné et al., 2002]. Figure 3 shows the evolution of the daily average floc size and the three shape descriptors considered here together with the SVI. Looking at the mean shape parameter values during the onset of the severe bulking problem (around the 55th day), a change in all three shape descriptors can be remarked. First, there was a decrease in roundness on day 59 (Figure 3(a)), and a simultaneous increase in reduced radius of gyration (Figure 3(c)). Hereafter, both parameters stabilized at their lower and higher value, respectively. This means that the activated sludge flocs evolved from having a somewhat circular shape to a more elongated one. This phenomenon may be explained by the high abundance of filaments, causing a more stretched type of floc to be formed, about four days after the bulking event started. Secondly, there was a change in the form factor of flocs, but unlike the other two parameters, the form factor started decreasing before filamentous bulking occurred (Figure 3(b)). Together with an SVI increase, which started on day 45, the flocs evolved slowly from smooth shapes to rougher ones, and finally the form factor stabilized at the smaller value from the 59th day on. On the other hand, the mean floc equivalent diameter D_{eq} somewhat decreased on the 19th day reaching its lowest value on the 42nd day and then increased thereafter as illustrated in Figure 3(d). This trend can be interpreted as a deflocculation followed by a flocculation of the floc forming biomass.

3.2 Modeling as a function of image analysis information

ARX models were identified so as to model the SVI evolution (i.e., the model output) based on information gathered with the image analysis procedure (i.e., the model input(s)). Since we are interested in modeling the *onset* of the bulking process, the first 70 data points for both input and output were taken instead of the entire 103 days.

3.2.1 ARX models

ARX models relate the current output $y(t)$ to a finite number of past outputs $y(t-k)$ and inputs $u(t-k)$.

$$y(t) + a_1 y(t-1) + (\cdots) + a_{na} y(t-na) = b_1 u(t-nk) + b_2 u(t-nk-1) + (\cdots) + b_{nb} u(t-nk-nb+1) + e(t) \quad (5)$$

with $y(t)$ equal to the output response at discrete time t, $u(t)$ the input at discrete time t, na the number of poles, nb the number of zeros, nk the pure time-delay (the dead-time) in the system and $e(t)$ a white noise signal. a_i and b_j are model parameters, with i = 1 ... na and j = 1 ... nb. The model structure is entirely defined by the three integers na, nb, and nk.

ARX models are identified by means of the ARX command in the System Identification Toolbox $5.0.1$ in MATLAB (The Mathworks, Inc., Natick), which allows to specify a specific *focus* during identification of the models. Three different options are available, i.e., a focus on *prediction*, *simulation* or *stability*. *Prediction* means that the model is determined by minimizing the prediction errors. With focus on *simulation*, the model approximation is such that the model will produce as good simulations as possible, when applied to inputs with the same spectra as used for the estimation. A stable model is guaranteed. Finally, a *stability* focus implies that the algorithm is modified so that a stable model is guaranteed, but the weighting still corresponds to prediction.

3.2.2 Optimization criteria

The criteria to be maximized are the *R-squared* (R^2) and *R-squared adjusted* (R^2_{adj}) values, both of which are often expressed in percent (Equations (6) and (7)). The values obtained from both criteria in this work reflect the percentage of output

variation explained by the model (i.e., $y_h(t)$). The R^2_{adj} criterion differs from the R^2 value in that it takes into account both the number of data points N and the model parameters (degrees of freedom) DF.

$$R^2 = 100 \cdot \Big(1 - \frac{\sum_{t=1}^{N}(y(t) - y_h(t))^2}{\sum_{t=1}^{N}(y(t) - mean(y(t)))^2}\Big) \quad (6)$$

$$R^2_{adj} = 100 \cdot \Big(1 - (N-1) \frac{\sum_{t=1}^{N}(y(t) - y_h(t))^2}{(N - DF) \cdot \sum_{t=1}^{N}(y(t) - mean(y(t)))^2}\Big) \quad (7)$$

with $y(t)$ the measured output at discrete time t, $y_h(t)$ the model output at discrete time t. Both performance indexes are used to evaluate the adequacy of the model produced.

From Figure 2(b) a very strong correlation between the SVI value and the total filament length per image could be noticed. Therefore, this input will be considered as single input first. Afterwards, other inputs (i.e., the equivalent diameter and the mean floc shape descriptors) will be taken into account as well.

3.2.3 Single input models

Optimal combinations in the range of 1 to 20 for the number of poles (na) and zeros (nb) (with nb smaller than or equal to na) at fixed delays nk of 0, 1 and 2 were sought. The results when the total filament length per image (F) is used as the ARX model input are shown in Table 1 with emphasis on *prediction*, *simulation* and *stability* for both criteria of R^2 and R^2_{adj}. Since a focus on *prediction* does not guarantee model stability, an additional stability check is performed. The optimal model without this stability check is shown on the 1st row while the 2nd row shows the usually less optimal (with respect to the criterion) but stable model. The values of na, nb, and nk with emphasis on *simulation* and *stabilization* are similar whether the system stability is explicitly accounted for or not.

Focus on	na	nb	nk	R^2 (%)	na	nb	nk	R^2_{adj} (%)
prediction	6	1	0	81.50	5	1	0	79.89
	6	1	0	81.50	5	1	0	79.89
	2	2	1	-2.94*10^4	2	1	1	-3.03*10^4
	-	-	1	-	-	-	1	-
	2	1	2	-4.11*10^3	2	1	2	-4.24*10^3
	-	-	2	-	-	-	2	-
Simulation	9	4	0	90.33	9	4	0	88.30
	11	9	1	83.16	11	9	1	76.76
	11	8	2	76.21	8	8	2	68.46
Stability	16	13	0	91.30	16	13	0	85.37
	16	16	1	77.47	6	3	1	71.09
	16	13	2	70.90	7	6	2	56.61

Table 1: Optimal ARX models at fixed nk values of 0, 1 or 2 and 1 model input, F.

As illustrated in Table 1, the optimal model with 0 delay (nk equal to zero) and focus on *prediction* is also a stable one while for higher delay values (nk equal to one or two), no stable models can be found. The 10th row, summarizes the ARX models (with emphasis on *stability*) which correspond to the highest value of R^2 (91.30%) and R^2_{adj} (85.37%), respectively. Figure 4 is oriented to reveal the performance of the 16th-order optimal ARX model (R^2=91.30%) with emphasis on *stability* in relation to the measured data.

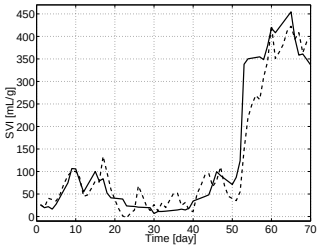


Figure 4: Measured (—) and modeled (- - -) SVI with the 16th-order optimal ARX model (R^2=91.30%) with emphasis on *stability* and 1 model input F.

3.2.4 Multiple inputs models

Table 2 summarizes the optimal combinations of the total filament length per image (F), the equivalent diameter (D_{eq}) and the mean floc shape descriptors, i.e., the roundness (R), the reduced radius of gyration (RG) and the form factor (FF) used as ARX model inputs with emphasis on *prediction* based on the R^2 criterion. Optimal combinations were sought in the order range of 1 to 15 and 1 to 9 for 2 or 3 model inputs, respectively. (It is worth mentioning that it was not possible to investigate higher ranges for the na values). Again, the 1st row shows the optimal models without stability check while the 2nd row shows the optimal stable ones.

na	nb					nk					R^2
	F	D_{eq}	FF	R	RG	F	D_{eq}	FF	R	RG	(%)
4	1	-	-	-	1	2	-	-	-	2	84.98
7	1	-	-	-	1	0	-	-	-	0	81.27
8	4	-	3	-	-	0	-	0	-	-	91.78
8	5	-	6	-	-	0	-	2	-	-	86.22
12	5	-	-	6	-	0	-	-	1	-	83.71
7	1	-	-	1	-	0	-	-	0	-	81.32
10	3	10	-	-	-	0	1	-	-	-	88.54
7	1	7	-	-	-	0	2	-	-	-	81.72
9	2	-	9	7	-	0	-	2	2	-	95.20
9	2	-	8	7	-	0	-	2	2	-	94.57
9	8	1	-	1	-	0	1	-	1	-	95.14
9	1	9	-	4	-	0	1	-	2	-	86.16
6	4	3	2	-	-	0	0	0	-	-	94.90
5	5	5	5	-	-	0	1	2	-	-	91.06

Table 2: Optimal ARX models with 2 or 3 model inputs and *prediction* focus.

The results in Table 2, the 9th row and the 10th row, both indicate that the 9th-order ARX models with emphasis on *prediction* and 3 model inputs (i.e., the total filament length per image (F), the mean floc form factor (FF) and the mean floc roundness (R)) yield the highest R^2 values of 95.20% and 94.57%, respectively.

Table 3 summarizes the optimal combinations of the total filament length per image (F), the equivalent diameter (D_{eq}) and

the mean floc shape descriptors, i.e., the roundness (R), the reduced radius of gyration (RG) and the form factor (FF) used as ARX model inputs with emphasis on *simulation* based on the R^2 criterion. Optimal combinations were sought in the order range of 1 to 15 and 1 to 9 for 2 or 3 model inputs, respectively.

na	nb					nk					R^2
	F	D_{eq}	FF	R	RG	F	D_{eq}	FF	R	RG	(%)
8	2	-	-	-	8	0	-	-	-	2	91.44
10	10	-	4	-	-	0	-	0	-	-	92.80
9	5	-	-	6	-	0	-	-	2	-	91.61
8	2	8	-	-	-	0	2	-	-	-	86.89
8	4	-	8	8	-	0	-	2	2	-	98.52
9	3	8	-	7	-	0	2	-	2	-	93.93
9	2	8	9	-	-	0	1	0	-	-	96.31

Table 3: Optimal ARX models with 2 or 3 model inputs and *simulation* focus.

The 5th row, shows that the 8th-order ARX model with emphasis on *simulation* and 3 model inputs, i.e., the total filament length per image (F), the mean floc form factor (FF) and the mean floc roundness (R) has the highest R^2 value of 98.52%. Figure 5 depicts the behavior of the 8th-order model with F, FF and R as the model inputs for the 70-day test period.

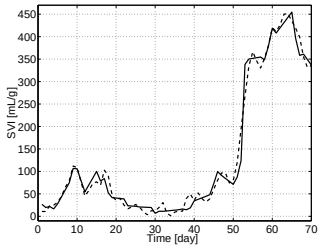


Figure 5: Measured (—) and modeled (- - -) SVI with the optimal 8th-order ARX model with emphasis on *simulation* and 3 model inputs (F, FF and R).

The results in Figure 5 and Table 3, confirm a strong agreement to experimental data obtained. The 8th-order ARX model output with 3 model inputs (F, FF and R) and R^2 of 98.52% lies substantially closer to the measured SVI data. Figure 6 shows the corresponding plot for the poles (denoted by $\times$) and the zeros (denoted by o).

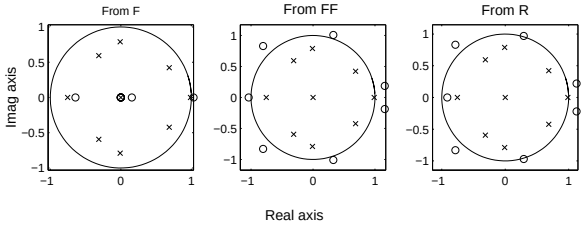


Figure 6: Location of the poles ($\times$) and zeros (o) for the optimal 8th-order ARX model with emphasis on *simulation* and 3 model inputs (F, FF and R).

Figure 6 illustrates that all the poles for the 8th-order ARX model with 3 model inputs (F, FF and R) with emphasis on *simulation* and R^2 of 98.52% lie in the unit circle implying a stable model output.

Table 4 summarizes the optimal combinations of the total filament length per image (F), the equivalent diameter (D_{eq}) and the mean floc shape descriptors, i.e., the roundness (R), the reduced radius of gyration (RG) and the form factor (FF) used as ARX model inputs with emphasis on *stability* based on the R^2 criterion. Optimal combinations were sought in the order range of 1 to 15 and 1 to 9 for 2 or 3 model inputs, respectively.

na	nb					nk					R^2
	F	D_{eq}	FF	R	RG	F	D_{eq}	FF	R	RG	(%)
8	8	-	-	-	2	0	-	-	-	0	87.07
14	13	-	3	-	-	0	-	0	-	-	90.96
8	8	-	-	2	-	0	-	-	2	-	87.54
15	15	14	-	-	-	0	0	-	-	-	92.01
9	3	-	9	8	-	0	-	2	2	-	95.59
9	9	6	-	1	-	0	0	-	0	-	92.44
8	3	8	6	-	-	0	2	2	-	-	93.30

Table 4: Optimal ARX models with 2 or 3 model inputs and *stability* focus.

It can be seen from the 5th row, that again a model with F, FF and R as the model inputs performs the best (R^2 = 95.59%). The model is now of order 9.

Table 5 summarizes the optimal combinations of the total filament length per image (F), the equivalent diameter (D_{eq}) and the mean floc shape descriptors, i.e., the roundness (R), the reduced radius of gyration (RG) and the form factor (FF) used as ARX model inputs with emphasis on *prediction* based on the R^2_{adj} criterion. Optimal combinations were sought in the order range of 1 to 15 and 1 to 9 for 2 or 3 model inputs, respectively.

na	nb					nk					R^2_{adj}
	F	D_{eq}	FF	R	RG	F	D_{eq}	FF	R	RG	(%)
4	1	-	-	-	1	2	-	-	-	2	83.81
6	1	-	-	-	1	0	-	-	-	2	79.08
8	4	-	3	-	-	0	-	0	-	-	89.69
8	3	-	6	-	-	0	-	2	-	-	81.64
6	1	-	-	1	-	0	-	-	0	-	79.04
6	1	-	-	1	-	0	-	-	0	-	79.04
10	3	10	-	-	-	0	1	-	-	-	83.17
5	1	1	-	-	-	0	1	-	-	-	79.57
9	2	-	9	7	-	0	-	2	2	-	92.30
8	2	-	8	7	-	0	-	2	2	-	91.67
9	8	1	-	1	-	0	1	-	1	-	93.42
3	3	2	-	1	-	0	0	-	2	-	84.04
6	4	3	2	-	-	0	0	0	-	-	93.60
5	3	4	5	-	-	0	1	2	-	-	88.30

Table 5: Optimal ARX models with 2 or 3 model inputs and *prediction* focus.

From Table 5, one can deduce that the 5th-order model with emphasis on *prediction* and 3 model inputs (i.e., the total filament length per image (F), the mean equivalent diameter (D_{eq}) and the mean form factor (FF)), gives to rise to the highest R^2_{adj} values of 93.60% (the 13th row).

Table 6 summarizes the optimal combinations of the total filament length per image (F), the equivalent diameter (D_{eq}) and the mean floc shape descriptors, i.e., the roundness (R), the reduced radius of gyration (RG) and the form factor (FF) used as ARX model inputs with emphasis on *simulation* based on the R^2_{adj} criterion. Optimal combinations were sought in the order range of 1 to 15 and 1 to 9 for 2 or 3 model inputs, respectively.

na	nb					nk					R^2_{adj} (%)
	F	D_{eq}	FF	R	RG	F	D_{eq}	FF	R	RG	
8	2	-	-	-	8	0	-	-	-	2	88.63
3	3	-	2	-	-	0	-	0	-	-	90.14
8	2	-	-	2	-	0	-	-	2	-	88.50
3	3	1	-	-	-	0	2	-	-	-	83.60
8	2	-	8	8	-	0	-	2	2	-	97.59
4	4	1	-	1	-	0	1	-	2	-	91.72
9	2	7	7	-	-	0	2	2	-	-	94.15

Table 6: Optimal ARX models with 2 or 3 model inputs and *simulation* focus.

It is observed in Table 6, 5th row, that the 8th-order model with emphasis on *simulation* and 3 model inputs (i.e., F, FF and R) generates the highest R^2_{adj} value of 97.59%.

Table 7 summarizes the optimal combinations of the total filament length per image (F), the equivalent diameter (D_{eq}) and the mean floc shape descriptors, i.e., the roundness (R), the reduced radius of gyration (RG) and the form factor (FF) used as ARX model inputs with emphasis on *stability* based on the R^2_{adj} criterion. Optimal combinations were sought in the order range of 1 to 15 and 1 to 9 for 2 or 3 model inputs, respectively.

na	nb					nk					R^2_{adj} (%)
	F	D_{eq}	FF	R	RG	F	D_{eq}	FF	R	RG	
8	8	-	-	-	1	0	-	-	-	2	82.90
8	8	-	4	-	-	0	-	2	-	-	87.20
8	8	-	-	1	-	0	-	-	2	-	83.74
14	13	1	-	-	-	0	0	-	-	-	84.17
9	3	-	9	8	-	0	-	2	2	-	92.57
8	8	7	-	1	-	0	0	-	0	-	88.93
8	3	7	5	-	-	0	2	2	-	-	89.83

Table 7: Optimal ARX models with 2 or 3 model inputs and *stability* focus.

The results in Table 7, 5th row, show that the 9th-order ARX model with emphasis on *stability* and 3 model inputs (i.e., F, FF and R) is the highest with an R^2_{adj} of 92.57%.

4 Conclusions

In the search for an early detection tool for filamentous bulking in activated sludge systems, a number of ARX model structures have been presented and tested to model the sludge volume index SVI (i.e., the output) as a function of image analysis information (i.e., the inputs). Based on the two performance quality criteria expressed in percent, i.e., R^2 and R^2_{adj}, multiple model inputs namely, 3 model inputs (i.e., the total filament length per image (F), the mean floc form factor (FF) and the mean floc roundness (R)) generated the relatively higher values irrespective of the emphasis taken into account. More to that, the 8th-order model with emphasis on *simulation* and 3 model inputs (i.e., F, FF and R) and an R^2 of 98.52% describes the sludge volume index SVI (i.e., the output) best thus far (Figure 5). It correctly reproduces the main dynamic characteristics the measured sludge volume index SVI (i.e., the output). The obtained models will be validated with new experimental data from ongoing experiments. Furthermore, the development of state space models is envisaged with the same inputs as exploited here. The results in this paper encourage further work on model development for monitoring and control purposes.

Acknowledgements

Work supported by Projects OT/99/24 and IDO/00/008 of the KULeuven Research Council, and the Belgian Program on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Ministers Office for Science, Technology and Culture. Ilse Smets is a postdoctoral fellow with the fund for Scientific Research Flanders. The scientific responsibility is assumed by its authors.

References

C. Cenens, K.P. Van Beurden, R. Jenné, and J.F. Van Impe. On the development of a novel image analysis technique to distinguish between flocs and filaments in activated sludge images. *Water Science and Technology*, 46(1-2):381–387, 2002.

M. da Motta, M.N. Pons, and N. Roche. Study of filamentous bacteria by image analysis and relation with settleability. *Water Science and Technology*, 46:363–369, 2002.

L. Eriksson and A.M. Hardin. Settling properties of activated sludge related to floc structure. *Water Science and Technology*, 16(10-11):55–68, 1984.

J.J. Ganczarczyk. Microbial aggregates in wastewater treatment. *Water Science and Technology*, 30(8):87–95, 1994.

C.A. Glasbey and G.W. Horgan. *Image analysis for the biological sciences*. John Wiley and Sons, New York (USA), 1995.

J. Houtmeyers, E. Van den Eynde, R. Poffé, and H. Verachtert. Relations between substrate feeding pattern and development of filamentous bacteria in activated sludge processes. Part I: Influence of process parameters. *European Journal of Applied Microbiology*, 9:63–77, 1980.

D. Jenkins, M. G. Richard, and G. T. Daigger. *Manual on the causes and control of activated sludge bulking and foaming*. 2nd edition. Lewis, London (UK), 1993.

R. Jenné, E.N. Banadda, N. Philips, and J. F. Van Impe. Image analysis as a monitoring tool for activated sludge properties in lab-scale installations. *Environmental Science and Health Part A (Accepted)*, 2003.

R. Jenné, C. Cenens, A.H. Geeraerd, and Van Impe J.F. Towards on-line quantification of flocs and filaments by image analysis. *Biotechnology Letters*, 24(11):931–935, 2002.

M.N. Pons, H. Vivier, J.F. Rémy, and J.A. Dodds. Morphological characterisation of yeast by image analysis. *Biotechnology and Bioengineering*, 42:1352–1359, 1993.

J.C. Russ. *Computer Assisted Microscopy: The Measurement and Analysis of Images*. Plenum Press, New York (USA), 1990.